# Iterative Methods for Approximating the Subdominant Modulus of an Eigenvalue of a Nonnegative Matrix

Moshe Haviv
*Department of Statistics*
*The Hebrew University*
*Jerusalem 91905, Israel*

Yakov Ritov
*Department of Statistics*
*The Hebrew University*
*Jerusalem 91905, Israel*

Uriel G. Rothblum*
*Faculty of Industrial Engineering and Management*
*Technion – Israel Institute of Technology*
*Haifa 32000, Israel*

Submitted by Hans Schneider

ABSTRACT

Two easily computable sequences of bounds on the subdominant modulus of an eigenvalue of a square nonnegative matrix are obtained. In particular it is shown that the sequences converge to the subdominant modulus. A sequence of bounds generated by a method of Brauer (1971) turns out to be a subsequence of one of our sequences. Thus, our results imply the convergence of Brauer's sequence.

## 1. INTRODUCTION

Hoffman [3], Brauer [2], and Rothblum and Tan [6] use the following method to obtain bounds on the subdominant modulus of an eigenvalue of an

*Research was supported by the Center for Absorption of Scientists, Ministry for Absorption and Immigration, The State of Israel.




52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

irreducible nonnegative matrix $B$ when its spectral radius and a corresponding positive eigenvector are given. First, from each column of the matrix, one deducts the maximal scalar multiple of the eigenvector so that nonnegativity is preserved. Then the spectral radius of the resulting matrix (which is easy to compute, since the corresponding eigenvector remains the original one) gives a desired bound. We apply this method on powers of a given matrix to obtain a sequence of bounds on the subdominant modulus of an eigenvalue of the original matrix and show that the resulting sequence converges to the subdominant modulus. Moreover, the convergence is monotonic on the subsequence corresponding to $B, B^2, B^4, B^8, B^{16}, \ldots$. This subsequence is shown to coincide with a sequence of bounds generated by an iterative method of Brauer [2]. Corresponding results are obtained when from each column of the matrix one deducts the minimal scalar multiple of the positive eigenvector so that the resulting matrix is nonpositive.

An extensive survey on bounds on the second largest modulus of an eigenvalue of a nonnegative matrix can be found in [6] where a unified method for generating bounds is given and is shown to produce or improve on most bounds that have appeared in the literature. The iterative methods developed here for generating bounds have two important virtues. First, they are very simple to implement computationally. Second, the generated sequence of bounds becomes tight (as it converges to the number one wishes to approximate).

We recall that the subdominant modulus of an eigenvalue of a corresponding irreducible, nonnegative matrix determines the geometric convergence rate of its normalized powers to a limit (e.g., [6]). Since our method requires the computation of the powers of the matrix, we are, in fact, computing the elements of a sequence in the process of obtaining bounds on its convergence rate. Moreover, taking powers of a matrix is expensive from a numerical point of view. Thus, we expect that computing the corresponding bounds for powers of a given matrix will not be efficient numerically. Of course, one might possibly be able to combine techniques from numerical analysis with our results to obtain efficient computational methods.

After introducing some notational conventions in Section 2, we state our main result in Section 3 and prove them in Section 5. The methods we use are illustrated on an example in Section 4.

## 2. NOTATION AND CONVENTIONS

We continue by introducing some notational conventions. The *spectrum* of an $S \times S$ matrix $A$ will be denoted $\sigma(A)$, and its *spectral radius* will be denoted $\rho(A)$; in particular $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$. Also, $\|A\|$ will denote the *$l_\infty$-operator norm* of $A$, i.e., $\|A\| = \max_{i=1,\dots,S} \sum_{j=1}^{S} |A_{ij}|$. A matrix $A$ is called *nonnegative* or *positive*, written $A \geqslant 0$ or $A \gg 0$, if all of its elements are *nonnegative* or *positive*, respectively. A matrix $A$ is called *semipositive*, written $A > 0$, if $A \geqslant 0$ and $A \neq 0$. These definitions will also apply to vectors in the natural way.

Let $B \in R^{S \times S}$ be nonnegative. The matrix $B$ is called *irreducible* if for some integer $N \geqslant 0$, $\sum_{j=0}^{N} B^j \gg 0$, and *irreducible and aperiodic* if for some integer $N \geqslant 0$, $B^N \gg 0$. The Perron-Frobenius theorem (e.g., [1]) assures that $\rho(B) \in \sigma(B)$ and that $B$ has semipositive left and right eigenvectors corresponding to $\rho(B)$. Moreover if $B$ is irreducible, then $B$ has positive left and right eigenvectors corresponding to $\rho(B)$. Next, we define the *subdominant modulus of an eigenvalue of* $B$, denoted $\xi(B)$, to be $\max\{|\lambda| : \lambda \in \sigma(B) \setminus \{\rho(B)\}\}$. This quantity is known to determine the convergence rate of iterative processes based on the matrix $B$ (see [6]). Thus, bounds on $\xi(B)$ determine bounds on the corresponding convergence rate. An extensive survey on the literature which examines such bounds can be found in [6].

The following lemma is key to our development. It shows how the spectrum of a square matrix is affected by certain transformations (see [2]).

LEMMA 1. *Let $A \in R^{S \times S}$, and let $w$ be a right eigenvector of $A$ corresponding to the eigenvalue $\lambda$. Then for every vector $a \in R^S$*

$$\sigma(A - wa^T) \supseteq (\sigma(A) \setminus \{\lambda\}) \cup \{\lambda - a^T w\}. \tag{1}$$

*Proof.* As $(A - wa^T)w = \lambda w - (a^T w)w = (\lambda - a^T w)w$, we have that $\lambda - a^T w \in \sigma(A - wa^T)$. Also, if $\gamma \neq \lambda$ is in $\sigma(A)$ and $z$ is a left eigenvector of $A$ corresponding to $\gamma$, then $\gamma z^T w = (z^T A)w = z^T(Aw) = \lambda z^T w$, implying that $z^T w = 0$. Hence, $z^T(A - wa^T) = \gamma z^T - (z^T w)a = \gamma z^T$, assuring that $\gamma \in \sigma(A - wa^T)$. ■

## 3. APPROXIMATING THE SUBDOMINANT MODULUS OF AN EIGENVALUE

Throughout the remainder of this paper we assume that $B$ is an $S \times S$ nonnegative matrix and that a positive right eigenvector $w$ of $B$ corresponding to $\rho(B)$ is given.[1] We next use Lemma 1 to obtain bounds on $\xi(B)$ in

[1] We note that when $B$ does not have a positive right eigenvector corresponding to $\rho(B)$, then necessarily $B$ is reducible. In this case the diagonal submatrices of $B$ corresponding to the classes of $B$ can be examined separately.

terms of $\rho(B)$ and $w$. The result first appeared in Hoffman [3] and was later obtained independently by Brauer [2]. It is included here for the sake of completeness.

LEMMA 2. *Let $b \in R^S$. Then*

$$\xi(B) \leqslant \rho(B - wb^T), \tag{2}$$

*and equality holds in* (2) *if $b^T w = \rho(B)$. Moreover, if $B - wb^T \geqslant 0$, then*

$$\xi(B) \leqslant \rho(B - wb^T) = \rho(B) - b^T w. \tag{3}$$

*Also, if $B - wb^T \leqslant 0$, then*

$$\xi(B) \leqslant \rho(B - wb^T) = b^T w - \rho(B). \tag{4}$$

*Proof.* The inequality (2) follows immediately from Lemma 1. Also, Lemma 1 shows that if $b^T w = \rho(B)$, then $\sigma(B - wb^T) = [\sigma(B) \setminus \rho(B)] \cup \{0\}$, immediately implying that in this case $\rho(B - wb^T) = \xi(B)$. We next observe that $(B - wb^T)w = Bw - (b^T w)w = [\rho(B) - b^T w]w$. Hence, if $B - wb^T \geqslant 0$, a standard result concerning nonnegative matrices (e.g., [1, Corollary 2.1.12]) implies that since $w \gg 0$, we have that $\rho(B - wb^T) = \rho(B) - b^T w$. Similarly, if $B - wb^T \leqslant 0$ then $wb^T - B \geqslant 0$; thus, as $(wb^T - B)w = [b^T w - \rho(B)]w$, we have that $\rho(B - wb^T) = \rho(wb^T - B) = b^T w - \rho(B)$. ∎

For $b \in R^S$, $B - wb^T \geqslant 0$ if and only if for $i, j = 1, \dots, S$, $w_i b_j \leqslant B_{ij}$. Thus, in this case, the tightest bound on $\xi(B)$ given by the right-hand side of (3) is obtained by selecting $b = \bar{b}(B, w)$, where

$$\bar{b}(B, w)_j = \min_{1 \leqslant i \leqslant S} w_i^{-1} B_{ij}, \qquad j = 1, \dots, S. \tag{5}$$

Under this selection of $b$, $\rho(B - wb^T) = \rho(B) - b^T w$ (see Lemma 2) will be denoted $\bar{\tau}(B, w)$, i.e.,

$$\bar{\tau}(B, w) = \rho(B) - \sum_{j=1}^{S} \Big( \min_{1 \leqslant i \leqslant S} w_i^{-1} B_{ij} \Big) w_j. \tag{6}$$

Similarly, the tightest bound on $\xi(B)$ given by the right-hand side of (4) when $B - wb^T \leqslant 0$ is obtained by selecting $b = \underline{b}(B, w)$, where

$$\underline{b}(B,w)_j = \max_{1 \leqslant i \leqslant S} w_i^{-1} B_{ij}, \qquad j = 1, \dots, S, \tag{7}$$

and the corresponding bound will be denoted $\underline{\tau}(B, w)$, i.e.,

$$\underline{\tau}(B,w) = \sum_{j=1}^{S} \Big( \max_{1 \leqslant i \leqslant S} w_i^{-1} B_{ij} \Big) w_j - \rho(B). \tag{8}$$

We note that $\bar{\tau}(B, w)$ and $\underline{\tau}(B, w)$ can be computed from $B$ and $w$ without the explicit use of $\rho(B)$. To compute $\bar{\tau}(B, w)$, determine $\bar{b}(B, w)$ from (5) and observe that

$$\left[ B - w\bar{b}(B,w)^T \right] w = \left[ \rho(B) - \bar{b}(B,w)^T w \right] w = \bar{\tau}(B,w) w. \tag{9}$$

Thus, $[B - w\bar{b}(B, w)^T]w$ is proportional to $w$, and the proportion coefficient is $\bar{\tau}(B, w)$. In particular, for $i = 1, \dots, S$

$$\bar{\tau}(B,w) = \frac{\left\{ \left[ B - w\bar{b}(B,w)^T \right] w \right\}_i}{w_i}. \tag{10}$$

Similarly, $[B - w\underline{b}(B, w)^T]w$ is proportional to $w$ and $-\underline{\tau}(B, w)$ is the proportion coefficient. Thus, for $i = 1, \dots, S$

$$\underline{\tau}(B,w) = -\frac{\left\{ \left[ B - w\underline{b}(B,w)^T \right] w \right\}_i}{w_i}. \tag{11}$$

We next state our main result. It asserts that by taking roots of the above bounds applied to corresponding powers of $B$, one gets a sequence of bounds on $\xi(B)$ which converges to $\xi(B)$. (Of course, $w$ is a right eigenvector of all powers of $B$ with respect to their spectral radius.) The proof is deferred to Section 5.

THEOREM 1. *For $k = 1, 2, \dots,$*

$$\xi(B) \leqslant \bar{\tau}(B^k, w)^{1/k}. \tag{12}$$

*Moreover, the limit of the sequence* $\{\bar{\tau}(B^k, w)^{l/k}\}$ *as* $k \to \infty$ *exists and*

$$\lim_{k \to \infty} \bar{\tau}(B^k, w)^{1/k} = \xi(B). \tag{13}$$

*Furthermore, the subsequence* $\{\bar{\tau}(B^{2^q}, w)^{1/2^q}\}$ *is monotonically decreasing. Finally, the same results hold with* $\underline{\tau}$ *replacing* $\bar{\tau}$.

We notice that (13) resembles a convergence of bounds on $\xi(B)$ which use powers of $B$ given in [6, Theorem 3.3].

We next demonstrate, by an example, that in general the sequence $\tau(B, w)^{1/k}$ is not monotonically decreasing (though Theorem 1 asserts that a specific subsequence of this sequence is). For $0 \leqslant \alpha \leqslant 1$, let

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1-\alpha & 0 & \alpha \end{pmatrix}.$$

Evidently, $\rho(P) = 1$, $Pe = e$ for $e = (1,1,1)^T$, and

$$P^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1-\alpha & 0 & \alpha \\ \alpha(1-\alpha) & 1-\alpha & \alpha^2 \end{pmatrix}$$

and

$$P^3 = \begin{pmatrix} 1-\alpha & 0 & \alpha \\ \alpha(1-\alpha) & 1-\alpha & \alpha^2 \\ \alpha^2(1-\alpha) & \alpha(1-\alpha) & 1-\alpha+\alpha^3 \end{pmatrix}.$$

So $\bar{b}(P, e) = (0,0,0)^T$, $\bar{b}(P^2, e) = (0,0,\alpha^2)^T$, and $\bar{b}(P^3, e) = (\alpha^2(1-\alpha), 0, \alpha^2)$. Thus, $\bar{\tau}(P^2, e) = 1$, $\bar{\tau}(P^2, e) = 1 - \alpha^2$, and $\bar{\tau}(P^3, e) = 1 - \alpha^2(1-\alpha) - \alpha^2 = (1-\alpha)(1+\alpha-\alpha^2)$. It follows that $\bar{\tau}(P^2, e)^{1/2} < \bar{\tau}(P^3, e)^{1/3}$ whenever $1/\sqrt{2} < \alpha \leqslant 1$.

Brauer [2] suggested the generation of a monotonically decreasing sequence of bounds on $\xi(B)$. To see how this sequence is determined, we define a sequence of $S \times S$ matrices $\{B(q)\}$ by letting $B(1) \equiv B$ and for $q = 1, 2, \ldots$

$$B(q+1) \equiv \left[B(q) - w\bar{b}(B(q), w)^T\right]^2. \tag{14}$$

Of course, in order to use this recursive formula one has to argue that for

$q=1,2,\ldots,$ $B(q)\geqslant 0$ and that $w$ is a (positive) right eigenvector of $B(q)$ corresponding to $\rho(B(q))$. This assertion is trite for $q=1$. Also, if it holds for $q$, then the definition of $\bar{b}(\cdot,\cdot)$ assures that $B(q)-w\bar{b}(B(q),w)^T$ is non-negative and therefore, by (14) so is $B(q+1)$. Also, by (9),

$$B(q+1)w=\left[B(q)-w\bar{b}(B(q),w)^T\right]^2 w=\left[\bar{\tau}(B(q),w)\right]^2 w,$$

and therefore a standard result concerning nonnegative matrices (e.g., [1, Corollary 2.1.12]) implies that $\rho(\mathrm{B}(\mathrm{q}+1))=\bar{\tau}(\mathrm{B}(\mathrm{q}),w)^2$ and that $w$ is a (positive) right eigenvector of $B(q+1)$ corresponding to $\rho(\mathrm{B}(\mathrm{q}+1))$. Brauer's sequence is the sequence $\bar{\tau}(\mathrm{B}(\mathrm{q}),w)^{1/2^q}$. Our second result asserts that this sequence is a subsequence of $\{\bar{\tau}(B^k,w)^{1/k}\}$. The proof is also deferred to Section 5.

THEOREM 2. *Let* $B(1)=B$ *and* $B(2),B(3),\ldots$ *be defined recursively by* (14). *Then for* $q=0,1,\ldots$

$$\bar{\tau}(B(q+1),w)=\bar{\tau}(B^{2^q},w)\leqslant\xi(B)^{2^q}. \tag{15}$$

*In particular, the sequence* $\{\bar{\tau}(B(q),w)^{1/2^q}\}$ *is decreasing and*

$$\lim_{q\to\infty}\bar{\tau}(B(q),w)^{1/2^q}=\xi(B). \tag{16}$$

We note without further details that one can get a representation of $\underline{\tau}(B^{2^q},w)$ similar to the one obtained in (15) for $\bar{\tau}(B^{2^q},w)$.

## 4. AN EXAMPLE

Let

$$B=\begin{pmatrix}4&3&3\\2&4&4\\3&4&3\end{pmatrix}.$$

As $Be=10e$ for $e=(1,1,1)^T$, we have that for $k=1,2,\ldots,$ $\rho(B^k)=10^k$, and $e$ is a corresponding right eigenvector. Straightforward computation yields the results summarized in Tables 1 and 2. Also, as (trivially)

$$\sigma\left(B-e\bar{b}(B,e)^T\right)=\left\{2,2^{-1}(1+\sqrt{5}),2^{-1}(1-\sqrt{5})\right\},$$

TABLE 1

| $k$ | $B^k$ | $\bar{b}(B^k,e)^T$ | $B^k - e\bar{b}(B^k,e)^T$ | $\bar{\tau}(B^k,e)^{1/k}$ |
|---|---|---|---|---|
| 1 | $\begin{pmatrix} 4 & 3 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 3 \end{pmatrix}$ | (2,3,3) | $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ | 2 |
| 2 | $\begin{pmatrix} 31 & 36 & 33 \\ 28 & 38 & 34 \\ 29 & 37 & 34 \end{pmatrix}$ | (28,36,33) | $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ | $3^{1/2} = 1.732$ |
| 4 | $\begin{pmatrix} 2926 & 3705 & 3369 \\ 2918 & 3710 & 3372 \\ 2921 & 3708 & 3371 \end{pmatrix}$ | (2918,3705,3369) | $\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 3 \\ 3 & 3 & 2 \end{pmatrix}$ | $8^{1/4} = 1.682$ |

TABLE 2

| $q$ | $B(q)$ | $\bar{b}(B(q),e)^T$ | $B(q) = e\bar{b}(B(q),e)^T$ | $\bar{\tau}(B(q),e)^{1/2^q}$ |
|---|---|---|---|---|
| 1 | $\begin{pmatrix} 4 & 3 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 3 \end{pmatrix}$ | (2,3,3) | $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ | 2 |
| 2 | $\begin{pmatrix} 4 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}$ | (1,0,0) | $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ | $3^{1/2} = 1.732$ |
| 3 | $\begin{pmatrix} 9 & 0 & 0 \\ 1 & 5 & 3 \\ 4 & 3 & 2 \end{pmatrix}$ | (1,0,0) | $\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 3 \\ 3 & 3 & 2 \end{pmatrix}$ | $8^{1/4} = 1.682$ |
| 4 | $\begin{pmatrix} 64 & 0 & 0 \\ 9 & 34 & 21 \\ 30 & 21 & 13 \end{pmatrix}$ | (9,0,0) | $\begin{pmatrix} 55 & 0 & 0 \\ 0 & 34 & 21 \\ 21 & 21 & 13 \end{pmatrix}$ | $55^{1/8} = 1.650$ |

we have from Lemma 2 that

$$\xi(B) = \xi\left(B - e\bar{b}(B,e)^T\right) = 2^{-1}(1+\sqrt{5}) = 1.618.$$

## 5. PROOFS

We will only establish the parts of Theorem 1 that concern $\bar{\tau}$, as the parts that concern $\underline{\tau}$ follow from similar arguments. Our proof will use two auxiliary lemmas. The first summarizes some properties of $\bar{b}(\cdot,\cdot)$ and $\bar{\tau}(\cdot,\cdot)$.

LEMMA 3. *For* $b \in R^S$,

$$B - wb^T \geqslant 0 \quad \textit{if and only if} \quad b \leqslant \bar{b}(B, w). \tag{17}$$

*In particular for b satisfying the equivalent conditions of* (17),

$$\rho(B) - b^T w \geqslant \bar{\tau}(B, w) \tag{18}$$

*with equality holding if and only if* $b = \bar{b}(B, w)$. *Also,*

$$\xi(B) \leqslant \bar{\tau}(B, w) \tag{19}$$

*and*

$$\left[B - w\bar{b}(B, w)^T\right] w = \bar{\tau}(B, w) w. \tag{20}$$

*Finally, for* $r, s = 1, 2, \ldots,$

$$\bar{\tau}(B^{r+s}, w) \leqslant \left[\bar{\tau}(B^r, w)\right]\left[\bar{\tau}(B^s, w)\right]. \tag{21}$$

*Proof.* The definition of $\bar{b}(B, w)$ directly implies the equivalence of the conditions in (17). Also, if $b \leqslant \bar{b}(B, w)$, then the positivity of $w$ implies that $b^T w \leqslant \bar{b}(B, w)^T w$ with equality holding if and only if $b = \bar{b}(B, w)$, i.e., $\rho(B) - b^T w \geqslant \rho(B) - \bar{b}(B, w)^T w = \bar{\tau}(B, w)$ with equality holding if and only if $b = \bar{b}(B, w)$. Next, as $B - w\bar{b}(B, w)^T \geqslant 0$, Lemma 2 implies (19). Also, trivially,

$$\left[B - w\bar{b}(B, w)^T\right] w = \rho(B) w - \left[\bar{b}(B, w)^T w\right] w = \bar{\tau}(B, w) w,$$

establishing (20). Next, let $b^k \equiv \bar{b}(B^k, w)$ for $k = 1, 2, \ldots,$ and let $\rho \equiv \rho(B)$. As

$$B^{r+s} - w\left[(b^r)^T B^s + \rho^r (b^s)^T - (b^r)^T w (b^s)^T\right]$$

$$= \left[B^r - w(b^r)^T\right]\left[B^s - w(b^s)^T\right] \geqslant 0,$$

we have from (18) (applied to $B^{r+s}$) that

$$\begin{aligned}\bar{\tau}(B^{r+s}, w) &\leqslant \rho(B^{r+s}) - \left[(b^r)^T B^s + \rho^r (b^s)^T - (b^r)^T w (b^s)^T\right] w \\ &= \rho^{r+s} - (b^r)^T \rho^s w - \rho^r (b^s)^T w + (b^r)^T w (b^s)^T w \\ &= \left[\rho^r - (b^r)^T w\right]\left[\rho^s - (b^s)^T w\right] = \bar{\tau}(B^r, w)\bar{\tau}(B^s, w),\end{aligned}$$

establishing (21). ∎

By the Perron-Frobenius theorem (see Section 2), $B$ has a semipositive eigenvector corresponding to its spectral radius. Throughout the rest of this section let $u$ be such an eigenvector satisfying the normalization condition $u^T w = 1$ (this is possible because $w \gg 0$).

We next consider bounds on $\xi(B)$ by applying Lemma 2 to vectors which are proportional to $u$. We will see that these bounds are never better than $\tau(B, w)$. However, we will use them to establish the desired convergence properties stated in Theorem 1.

For $\beta \in R$, $B - w(\beta u^T) \geqslant 0$ if and only if for all $i, j = 1, \ldots, S$, $\beta w_i u_j \leqslant B_{ij}$. Thus, the tightest bound on $\xi(B)$ given by the right-hand side of (3) for vectors $b$ which are proportional to $u$ is obtained by selecting $b = \bar{\beta}(B, w, u)u$, where

$$\bar{\beta}(B, w, u) = \min_{\substack{1 \leqslant i, j \leqslant S \\ u_j > 0}} w_i^{-1} u_j^{-1} B_{ij}. \tag{22}$$

Under this selection of $b$, $\rho(B - wb^T) = \rho(B) - b^T w$ (see Lemma 2) will be denoted $\bar{\mu}(B, w, u)$, i.e.,

$$\bar{\mu}(B, w, u) = \rho(B) - \bar{\beta}(B, w, u) u^T w = \rho(B) - \bar{\beta}(B, w, u). \tag{23}$$

The following lemma summarizes some properties of $\bar{\beta}(B, w, u)$ and $\bar{\mu}(B, w, u)$.

LEMMA 4. *For $\beta \in R$,*

$$B - \beta w u^T \geqslant 0 \quad \textit{if and only if} \quad \beta \leqslant \bar{\beta}(B, w, u). \tag{24}$$

*In particular, for $\beta$ satisfying the equivalent conditions of* (24),

$$\rho(B) - \beta \geqslant \bar{\mu}(B, w, u). \tag{25}$$

*Also,*

$$\xi(B) \leqslant \bar{\tau}(B, w) \leqslant \bar{\mu}(B, w, u) \tag{26}$$

*and*

$$[B - wu^T\bar{\beta}(B, w, u)]w = \bar{\mu}(B, w, u)w. \tag{27}$$

*Finally, for* $r, s = 1, 2, \ldots$

$$\bar{\mu}(B^{r+s}, w, u) \leqslant [\bar{\mu}(B^r, w, u)][\bar{\mu}(B^s, w, u)]. \tag{28}$$

*Proof.* The definition of $\bar{\beta}(B, , u)$ directly implies the equivalence of the conditions in (24). Also, if $\beta \leqslant \bar{\beta}(B, w, u)$, then $\rho(B) - \beta \geqslant \rho(B) - \beta u^T w \geqslant \rho(B) - \bar{\beta}(B, w, u) = \bar{\mu}(B, w, u)$ establishing (25). Next, as $B - \bar{\beta}(B, w, u)wu^T \geqslant 0$, Lemma 3 implies that $\tau(B, w) \leqslant \rho(B) - \bar{\beta}(B, w, u)u^T w = \bar{\mu}(B, w, u)$, establishing the second inequality in (26). The first inequality in (26) was established in Lemma 3. Next, clearly,

$$[B - wu^T\bar{\beta}(B, w, u)]w = [\rho(B) - \bar{\beta}(B, w, u)]w = \bar{\mu}(B, w, u)w,$$

establishing (27). Finally, let $\beta_k = \bar{\beta}(B^k, w, u)$ for $k = 1, 2, \ldots$, and let $\rho \equiv \rho(B)$. As

$$B^{r+s} - wu^T(\beta_r\rho^s + \beta_s\rho^r - \beta_r\beta_s) = [B^r - \beta_r wu^T][B^s - \beta_s wu^T] \geqslant 0,$$

we have from (25) (applied to $B^{r+s}$) that

$$\begin{aligned}\bar{\mu}(B, w, u) &\leqslant \rho(B^{r+s}) - \beta_r\rho^s - \beta_s\rho^r + \beta_r\beta_s \\ &= (\rho^r - \beta_r)(\rho^s - \beta_s) = [\bar{\mu}(B^r, w, u)][\bar{\mu}(B^s, w, u)],\end{aligned}$$

establishing (28). ∎

We are now ready for the proof of Theorem 1.

*Proof of Theorem 1.* Applying (26) to $B^k$, we have that

$$\xi(B)^k = \xi(B^k) \leqslant \bar{\tau}(B^k, w) \leqslant \bar{\mu}(B^k, w, u),$$

establishing (12). In order to establish the asserted convergences in (13), it remains to show that $\limsup_{k\to\infty}\bar{\mu}(B^k, w, u)^{1/k} = \xi(B)$. Let $\rho \equiv \rho(B)$, $\bar{\beta} \equiv \bar{\beta}(B, w, u)$, and $C \equiv B - \rho w u^T$. As $u^T \rho w = \rho$, we have from Lemma 2 that $\rho(C) = \xi(B)$. Also, as $Bw = \rho w$ and $u^T B = \rho u^T$, a simple inductive argument shows that $C^k = B^k - \rho^k w u^T$.

Let $\alpha \equiv \min\{w_i u_j : i, j = 1,\ldots, S,\ u_j > 0\}$. Evidently, for $i, j = 1,\ldots, S$, if $u_j > 0$, then

$$-(C^k)_{ij} \leqslant \|C^k\| \leqslant \alpha^{-1} w_i u_j \|C^k\| = \left(\alpha^{-1}\|C^k\| w u^T\right)_{ij},$$

and if $u_j = 0$ then $(wu^T)_{ij} = w_i u_j = 0$, assuring that

$$-(C^k)_{ij} = -(B^k)_{ij} \leqslant 0 = \left(\alpha^{-1}\|C^k\| w u^T\right)_{ij}.$$

It follows that

$$B^k - \left(\rho^k - \alpha^{-1}\|C^k\|\right) w u^T = C^k + \alpha^{-1}\|C^k\| w u^T \geqslant 0,$$

and therefore, by Lemma 4,

$$\bar{\mu}(B^k, w, u) \leqslant \rho(B^k) - \left(\rho^k - \alpha^{-1}\|C^k\|\right) = \alpha^{-1}\|C^k\|. \tag{29}$$

By the spectral-radius formula (e.g., [5, p. 98]), the limit of $\|C^k\|^{1/k}$ exists and equals $\rho(C) = \xi(B)$. Thus, we conclude from (29) that

$$\begin{aligned}\limsup_{k\to\infty} \bar{\mu}(B^k, w, u)^{1/k} &\leqslant \limsup_{k\to\infty}\left(\alpha^{-1}\|C^k\|\right)^{1/k} \\ &= \limsup_{k\to\infty}\|C^k\|^{1/k} = \rho(C) = \xi(B),\end{aligned}$$

completing the proof of (13). Finally, observe that (21) implies that for $k = 1, 2, \ldots,$ $\bar{\tau}(B^{2k}, w) \leqslant \bar{\tau}(B^k, w)^2$ and therefore $\bar{\tau}(B^{2k}, w)^{1/2k} \leqslant \bar{\tau}(B^k, w)^{1/k}$. This immediately implies that the sequence $\{\bar{\tau}(B^{2^q}, w)^{1/2^q}\}$ is monotonically decreasing. ■

We note that the established convergence of the sequences $\{\bar{\tau}(B^k, w)^{1/k}\}$ and $\{\bar{\mu}(B^k, w, u)^{1/k}\}$ also follows directly from (21) and (28), respectively, and the well-known fact that a sequence $\{f(k)\}$ satisfying the functional inequality $f(r+s) \leqslant f(r)f(s)$ converges (e.g., [4, p. 27]).

We next establish Theorem 2.

*Proof of Theorem* 2. We first show that for $q = 1,2,\ldots,$ there exists a vector $b(q) \in R^S$ such that

$$B(q) = B^{2^q} - wb(q)^T. \tag{30}$$

This result is trivial for $q = 1$ [with $b(1) = 0$]. Next assume that it holds for $q$, and consider $q + 1$. By (14) and the induction hypothesis

$$\begin{aligned} B(q+1) &= \left[B(q) - w\bar{b}(B(q), w)^T\right]^{2^q} \\ &= \left\{B^{2^q} - w\left[b(q)^T + \bar{b}(B(q), w)^T\right]\right\}^2 \end{aligned}$$

Thus, with $c(q) \equiv b(q) + \bar{b}(B(q), w)$,

$$B(q+1) = B^{2^{q+1}} - B^{2^q}wc(q)^T - wc(q)^T B^{2^q} + wc(q)^T wc(q)^T.$$

As $B^{2^q}w = \rho(B)^{2^q}w$, we conclude that $B(q+1) = B^{2^{q+1}} - wb(q+1)^T$, where

$$b(q+1)^T = \rho(B)^{2^q}c(q)^T + c(q)^T B^{2^q} - c(q)^T wc(q)^T,$$

thereby completing the inductive proof of (30).

Next, let $q = 1,2,\ldots$ be given. As $B(q) \geqslant 0$ and $w$ is a (positive) right eigenvector of $B(q)$ corresponding to $\rho(B(q))$, it follows immediately from (30) and the definition of $\bar{b}(\cdot, w)$ that $\bar{b}(B(q), w) = \bar{b}(B^{2^q}, w) - b(q)$ and therefore

$$\begin{aligned} \bar{\tau}(B(q), w) &= \rho\left(B(q) - w\bar{b}(B(q), w)^T\right) \\ &= \rho\left(B^{2^q} - w\left[b(q) + \bar{b}(B(q), w)\right]^T\right) \\ &= \rho\left(B^{2^q} - w\bar{b}(B^{2^q}, w)\right) = \bar{\tau}(B^{2^q}, w), \end{aligned}$$

establishing the equality in (15). Next, the inequality in (15) follows immediately by applying (19) to $B^{2^q}$ and using the fact that $\xi(B^{2^q}) = \xi(B)^{2^q}$. Finally, the monotonicity and convergence of the sequence $\{\bar{\tau}(B(q), w)^{1/2^q}\}$ follows directly from (15) and the results of Theorem 1. ∎

## 6. COMPUTATION

For $q = 1,2,\ldots,$ $\bar{\tau}(B^{2^q}, w)^{1/2^q} = \bar{\tau}(B(q), w)^{1/2^q}$ can be computed in two ways. The first amounts to computing

$$B^{2^q},\ \bar{b}(B^{2^q}, w),\ \rho\big(B^{2^q} - w\bar{b}(B^{2^q}, w)^T\big),$$

and the $2^q$th root of the last expression, while the second amounts to computing

$$B(q),\ \bar{b}(B(q), w),\ \rho\big(B(q) - w\bar{b}(B(q), w)^T\big),$$

and the $2^q$th root of the last expression. The successive computation of $B(q)$ requires $q$ squarings of $S \times S$ matrices and the determination of $\bar{b}(B(1), w), \ldots, \bar{b}(B(q-1), w)$, while computing $B^{2^q}$ requires only $q$ squarings of the $S \times S$ matrices $B, B^2, \ldots, B^{2^{q-1}}$. On the other hand, the example in Section 4 demonstrates that for $q = 1, 2, \ldots,$ $B(q) - w\bar{b}(B(q), w)^T$ has more zero elements than does $B^{2^q}$, and therefore squaring $B(q) - w\bar{b}(B(q), w)^T$ requires fewer arithmetic operations than squaring $B^{2^q}$. This advantage is noticeable only for small matrices. Moreover, when a computer is used, remembering the locations of the zero elements will offset the above advantage. We also note that *a priori* one cannot know whether $\bar{\tau}(B^{2^q}, w)$ or $\underline{\tau}(B^{2^q}, w)$ is a better bound. Thus, it makes sense to compute both.

We note that in order to avoid large numbers one can normalize the matrix $B$ and consider $\rho(B)^{-1}B$. Of course, $\xi(\rho(B)^{-1}B) = \rho(B)^{-1}\xi(B)$. Also, as $\xi(B^T) = \xi(B)$, one can use the methods of this paper on $B^T$ to get bounds on $\xi(B)$, which are usually different from those derived for $B$ directly. Another method to derive bounds on $\xi(B)$ is to scale the matrix $B$ by replacing it with $D^{-1}BD$, where $D$ is an $S \times S$ diagonal matrix having positive diagonal elements. As $\sigma(B) = \sigma(D^{-1}BD)$ we have that $\xi(B) = \xi(D^{-1}BD)$. The next lemma shows that if the methods of this paper are applied to $D^{-1}BD$ with its positive right eigenvector $D^{-1}w$, the resulting bounds are invariant under the choice of $D$. This is in contrast with bounds on $\xi(B)$ given in [6], where different scalings can result different bounds.

LEMMA 5. *Let $D \in R^{S \times S}$ be a diagonal matrix having positive diagonal elements. Then $\bar{\tau}(D^{-1}BD, D^{-1}w) = \bar{\tau}(B, w)$ and $\underline{\tau}(D^{-1}BD, D^{-1}w) = \underline{\tau}(B, w)$.*

*Proof.* Evidently, for $a \in R^S$, $B - wa^T \geqslant 0$ if and only if $D^{-1}BD - D^{-1}w(Da)^T \geqslant 0$. Thus, it follows from two applications of Lemma 3 that $D\bar{b}(B, w) = \bar{b}(D^{-1}BD, D^{-1}w)$, implying that $\bar{\tau}(D^{-1}BD, D^{-1}w) = \rho(D^{-1}BD) - \bar{b}(D^{-1}BD, D^{-1}w)^T D^{-1}w = \rho(B) - \bar{b}(B, w)^T DD^{-1}w = \bar{\tau}(B, w)$. Similar arguments establish that $\underline{\tau}(D^{-1}BD, D^{-1}w) = \underline{\tau}(B, w)$. ∎

If $D$ is diagonal matrix whose diagonal elements are $w_1,\dots,w_S$, we have that $D^{-1}w = e = (1,\dots,1)^T$ is a positive right eigenvector of $D^{-1}BD$. In this case the computation of $\bar{b}(D^{-1}BD, D^{-1}w) = \bar{b}(D^{-1}BD, e)$ by (5) and that of $\bar{\tau}(D^{-1}BD, e)$ by (6) simplify.

Looking at the sequence $\{\tau(B(q), w)^{1/2^q}\}$ generated by Brauer's method, we see that an improvement of the bounds from iteration $q$ to $q+1$ occurs only if $B(q+1)$ has a positive column. If this is not the case, then $\bar{b}(B(q), w) = 0$ and $B(q+1) = B(q)^2$. It follows from Seneta [7, p. 58] that if the bounds do not improve during $\lceil \log_2(S^2 - 2S + 2) \rceil$ successive iterations starting with iteration $q$, then $B(q)$ is either reducible or periodic, and therefore for all future iterations $B(\cdot)$ will have no positive column and the bounds will not improve. We conclude from Theorem 2 that if $\bar{b}(B^{2^q}, w) = 0$ for $\lceil \log_2(S^2 - 2S + 2) \rceil$ successive iterations, then no further improvements of the bounds $\tau(B^{2^q}, w)$ will occur, and therefore this situation indicates finite convergence [to $\xi(B)$].

## REFERENCES


1 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.

2 A. Brauer, A remark on the paper of H. H. Schaefer: Abschätzung der nichttrivalen Eigenwerte stochastischer Matrizen, *Numer. Math.* 17:163–165 (1971).

3 A. J. Hoffman, Three observations on nonnegative matrices, *J. Res. Nat. Bur. Standards — B. Math. and Math. Phys.* 71B:39–41 (1967).

4 T. Kato, *Perturbation Theory for Linear Operators*, Springer, New York, 1966.

5 B. V. Limaye, *Functional Analysis*, Wiley Eastern, 1981.

6 U. G. Rothblum and C. Tan, Upper bounds on the maximum modulus of the subdominant eigenvalue of nonnegative matrices, *Linear Algebra Appl.* 66:45–86 (1985).

7 E. Seneta, *Non-negative Matrices and Markov Chains*, Springer, New York, 1981.


*Received 16 January 1986; revised 20 January 1986*